Digital Content Strategies*

Daniel Halbheer, Florian Stahl, Oded Koenigsberg, and Donald R. Lehmann†

February 2013

Abstract

This paper studies content strategies for online publishers of digital information goods. It examines sampling strategies and compares their performance to paid content and free content strategies. A sampling strategy, where some of the content is offered for free and consumers are charged for access to the rest, is known as a "metered model" in the newspaper industry. We analyze optimal decisions concerning the size of the sample and the price of the paid content when sampling serves the dual purpose of disclosing content quality and generating advertising revenue. We show in a reduced-form model how the publisher's optimal ratio of advertising revenue to sales revenue is linked to characteristics of both the content market and the advertising market. We assume that consumers learn about content quality from the free samples in a Bayesian fashion. Surprisingly, we find that it can be optimal for the publisher to generate advertising revenue by offering free samples even when sampling reduces both prior quality expectations and content demand. In addition, we show that it can be optimal for the publisher to refrain from revealing quality through free samples when advertising effectiveness is low and content quality is high.

Keywords: Information Goods, Sampling, Content Pricing, Advertising, Dorfman-Steiner Condition

---

*We are grateful to the guest editor, Marnik Dekimpe, and three anonymous referees for helpful comments. We also thank Erwin Amann, Asim Ansari, Jean-Pierre Dubé, Anthony Dukes, Jacob Goldenberg, Avi Goldfarb, Raju Hornis, Ulrich Kaiser, Anja Lambrecht, Catherine Tucker and seminar participants at the GEABA 2012 (Graz), the Marketing in Israel Conference 2011 (Tel Aviv), the INFORMS Marketing Science Conference 2011 (Houston), the University of Hamburg, the University of Passau, the University of Tilburg, and the University of Zurich for helpful comments and suggestions. Daniel Halbheer gratefully acknowledges support from the Swiss National Science Foundation through grant PA00P1-129097 and thanks the Department of Economics at the University of Virginia for its hospitality while some of this research was being undertaken. A previous version of this paper circulated under the title "Sampling Strategies for Information Goods."

†Daniel Halbheer (corresponding author) and Florian Stahl: University of Zurich, Department of Business Administration, Plattenstrasse 14, CH-8032 Zurich, Switzerland (e-mail: daniel.halbheer@uzh.ch and florian.stahl@uzh.ch, respectively); Oded Koenigsberg: London Business School, Regent's Park, London, NW1 4SA, United Kingdom (e-mail: okoenigsberg@london.edu); Donald R. Lehmann: Columbia Business School, Uris Hall, 3022 Broadway, New York, NY 10027, United States (e-mail: drl2@columbia.edu).

1 Introduction

Digital information goods have been available on the Internet for almost twenty years. During that time, publishers have developed different strategies to distribute content. Some publishers provide all their information for free, while some charge consumers for access to their content. Other publishers employ a hybrid business model, giving away a portion of their content to consumers for free and charging for access to the rest of their content. Offering free content samples allows publishers to both disclose their content quality and to generate revenues from advertisements shown to online visitors. According to Alisa Bowen, general manager of *The Wall Street Journal Digital Network*, "working with advertisers to offer open houses has proven to be one of the most valuable and efficient ways to expose our premium content to new readers and potential subscribers" (*GlobeNewswire*, 2012). The main contribution of this paper is to provide a formal analysis of how publishers should choose between different digital content strategies.

Information goods are experience goods and offering free content samples is a way for publishers to disclose their product quality and allow consumers to have actual experience with the good before purchase (Shapiro and Varian, 1998). Digital information goods are particularly suitable for sampling because the costs of providing free samples are negligible and the publisher can include advertisements in the free samples to generate advertising revenues. These two features distinguish sampling of information goods from sampling perishable goods or durable goods.

Recently, hybrid business models where publishers set the size of the sample and consumers select the samples of their choice have emerged. A prominent example of this is the "metered model" in the newspaper industry, where publishers offer a number of articles for free and charge for access to the rest. Such "customer selected sampling" differs from the approach where the publisher chooses not only the sample size but also the sample content, which allows the firm to strategically manipulate the sample and creates an environment where customers are likely to discount the sample quality in estimating actual quality. A recent study by the *Newspaper Association of America* (2012) shows that 62% of the publishers employ a metered model, out of which 95% offer up to twenty free articles monthly. For example, the *New York Times* currently offers access to ten articles for free on its website each month. Advertising supported sampling is also employed by distributors of music such as *Spotify* or *Rhapsody*. Allowing consumers to choose which content to sample means publishers have no control over the content consumers actually sample. Taking this into account is important for publishers when setting the optimal sample size.

The business model where publishers set a sample size and let the consumers choose which content to sample differs from versioning or "freemium," where a firm selected low-end version is offered for free and consumers have to pay for access to the high-end version.[1] Such versioning of information goods is often observed in the software industry (see, for instance, Faugère and Tayi 2007; Cheng and Tang 2010). Customer selected sampling, in contrast, does not involve quality differentiation: Within the set limit, the publisher allows the consumers to sample any of its content for free.

This paper develops an analytical model to study optimal decisions concerning the size of the sample and the price of the paid content for online publishers of digital information goods when sampling serves the dual purpose of disclosing content quality and generating revenues from advertising. The publisher is assumed to receive revenues from content sales and from advertisements, which are included with the free content. Consumers have prior expectations about content quality, which they update in a Bayesian fashion through inspection of the free samples. The information transmitted through samples affects the consumers' posterior expectations about content quality, which in turn influence demand for the paid content (content demand). Taking the consumers' quality updating into account, the publisher faces a tradeoff between an expansion effect (through learning) and a cannibalization effect (through free offerings) on content demand induced by sampling. When the publisher makes its sampling and pricing decisions, it should take the two countervailing effects on content demand and on the advertising revenue into account. We assume that the publisher can either adopt a "sampling strategy," a pure "paid content strategy," or a pure "free content strategy."

We derive several results. *First*, we show, in a reduced-form model, how the publisher's optimal ratio of advertising revenue to sales revenue is determined by characteristics of both the content market and the advertising market. Specifically, the key determinants of the advertising-sales revenue ratio are the elasticities of expected content demand with respect to price and sample size, the price elasticity of advertising demand, and the elasticity of consumers' updated expectations with respect to the sample size. The latter plays a crucial role in the determining the ratio of advertising to sales revenue: When expectations are increasing in sample size, the ratio tends to be lower, whereas it tends to be higher if expectations are decreasing in sample size. This result arises because an increase in expectations mitigates or even compensates for the cannibalization effect, thus leading to a lower advertising-sales revenue ratio. If instead sampling reduces expectations, offering free samples reinforces the cannibalization effect, which in turn leads to a higher ratio of advertising revenue to sales revenue. Nevertheless, the

[1] Bhargava and Choudhary (2008) analyze optimal versioning of information goods.

publisher will engage in ad-supported sampling if the advertising price per impression is high enough.

Second, we characterize the publisher's optimal sample size and price decisions in a benchmark model where content quality is common knowledge. The optimal strategy is determined by the relationship between the advertising effectiveness and content quality. A paid content strategy is optimal for the publisher only if the effectiveness of advertising is sufficiently low. For intermediate levels of the advertising effectiveness, the publisher should employ a sampling strategy and generate revenues from both sales and advertising. Once advertising is sufficiently effective, the publisher should switch to a free content strategy. Thus, it may be optimal for the publisher to offer free content samples even if sampling cannibalizes content demand.

Third, we characterize the publisher's optimal sample size and price decisions when consumers learn about content quality through inspection of the free samples. Assuming that consumers are uncertain about content quality, sampling has a demand-enhancing effect when the elasticity of consumer's posterior expectations with respect to sample size exceeds the ratio of sampled to paid content. The optimal strategy is determined by the relationship between advertising effectiveness and the interplay between quality expectations and actual content quality. As in the benchmark model, employing a paid content strategy is optimal only if advertising effectiveness is sufficiently low compared to prior quality expectations, a sampling strategy is optimal for intermediate levels of advertising effectiveness, and the publisher should switch to a free content strategy once advertising is sufficiently effective compared to posterior quality expectations.

Our paper is related to two literature streams. The first stream is on media firm strategy in two-sided markets.[2] For instance, Kind et al. (2009) analyze how competition, captured by the number of media platforms and content differentiation between platforms, affects the composition of revenues from advertising and sales. Godes et al. (2009) investigate a similar question, focusing on competition between platforms in different media industries. Our paper examines optimal advertising supported content sampling and content pricing when the firm can derive revenue from content sales, advertising, or both. Papers that examine content sampling from different perspectives include Xiang and Soberman (2011) for preview provision and Chellappa and Shivendu (2005) for piracy-mitigating strategies, but neither consider the impact of sampling on advertising revenues. To the best of our knowledge, optimal content sampling when

[2]See Rysman (2009) for a general review of the two-sided markets literature. Anderson and Gabszewicz (2006) provide a canonical survey of media and advertising.

sampling impacts revenues from both content sales and online advertising has not been addressed by the literature.

This paper is also related to the broad literature on consumer learning about product attributes. In many markets, firms enable consumer learning through disclosing information about their products and services. Information can be disclosed in various ways: For instance, through informative advertising (see Anderson and Renault 2006, and Bagwell 2007 for a comprehensive survey), or product descriptions or third-party reviews (Sun 2011; Hotz and Xiao 2013). Another way for firms to disclose information is through sampling. Heiman et al. (2001) and Bawa and Shoemaker (2004) study how sampling affects demand and the evolution of market shares for consumer goods, while Boom (2009) and Wang and Zhang (2009) investigate sampling of information goods. However, when firms sample information goods, they only offer a portion of the good for free to avoid the "information paradox" (Akerlof, 1970). The consumers' inference from this portion about the product's attributes is most naturally modeled in a Bayesian framework. Bayesian learning processes based on product experience have been widely employed in the literature, for instance, by Erdem and Keane (1996), Ackerberg (2003), and Erdem et al. (2008), and we follow this approach here.

We organize the remainder of the paper as follows. Section 2 presents the general framework. Section 3 describes the model and the consumer's learning mechanism in particular. Section 4 characterizes optimal sampling and pricing decisions when consumers know content quality. Section 5 extends the analysis to the case of incomplete information and assumes that quality is initially the publisher's private information. Section 6 introduces two model extensions: the inclusion of advertisements in both the free articles and the paid content and competition among publishers. Conclusions and directions for future research are offered in Section 7. To facilitate exposition, we have relegated proofs to the Appendix.

2 General Framework

We now introduce the three main components of our modeling framework: the publisher, the consumers and the advertising market. We then define the strategies available to the publisher and derive the optimal advertising-sales revenue ratio.

We consider a publisher who offers a digital information good with content of size $N > 0$ through an online channel. Content size may be thought of as the number of chapters of a book or movie, the number of songs on an album, or the number of articles on a news platform. We assume that the publisher has constant unit costs $c \geq 0$

and fixed costs $F \geq 0$ to produce the content.[3] The cost to provide digital access per subscriber is $c_s \geq 0$ and the costs of providing free samples are normalized to zero. The qualities of the content parts are distributed on the quality spectrum $[0, \overline{V}]$, where $\overline{V}$ is the publisher's private information. We consider quality $\overline{V}$ as an outcome of a previous strategic decision and focus on the publisher's short-run pricing and sampling decisions. Thus, the publisher has two decision variables: the sample size $n \in [0, N]$ and the price p at which to sell the good.[4]

We consider a market with a unit measure of consumers that observe the publisher's sampling and pricing decisions. Consumers are uncertain about content quality. We assume that they update their prior expectations in a Bayesian fashion through inspection of the free samples and denote by $\tilde{V}(n)$ the consumers' expected posterior quality given the sample size n. The demand for paid content (content demand) depends on price p, sample size n, and $\tilde{V}(n)$. Specifically, we assume that the publisher's *expected* content demand is given by

$$D^E(p,n) \equiv D(p,n,\tilde{V}(n)). \tag{1}$$

This representation emphasizes that the sample size has both a direct effect on content demand and an indirect effect that operates through the impact of n on expected posterior quality $\tilde{V}(n)$.

We assume that content demand satisfies the following basic assumptions. First, we assume that $\frac{\partial D}{\partial p} < 0$, i.e. content demand depends negatively on price. Second, we impose that $\frac{\partial D}{\partial n} < 0$, so that a larger sample size has a *direct* negative effect on demand for the remaining content. Third, we require that $\frac{\partial D}{\partial \tilde{V}} > 0$, i.e. content demand depends positively on expected posterior quality. The overall effect of the sample size n on expected content demand is given by

$$\frac{\partial D^E}{\partial n} = \frac{\partial D}{\partial n} + \frac{\partial D}{\partial \tilde{V}} \tilde{V}'(n),$$

where term $\frac{\partial D}{\partial \tilde{V}} \tilde{V}'(n)$ captures the *indirect* effect of the sample size on expected content demand. It is not clear a priori how the sample size affects posterior expectations and hence $\frac{\partial D^E}{\partial n}$. If $\tilde{V}'(n) < 0$, sampling reduces posterior expectations and is thus *demand-reducing*. Note that even if $\tilde{V}'(n) > 0$, that is, if sampling increases posterior expectations, offering an additional sample may be demand-reducing if the direct effect domi-

[3] Throughout the analysis, we assume that the fixed cost do not exceed the product market profits. Hence they do not change the analysis and can therefore be omitted.

[4] The choice of (p,n) is not a multidimensional signal for quality as studied, for instance, by Wilson (1985) and Milgrom and Roberts (1984). In this strand of the literature, n is an advertising signal for quality. However, in our setting, the publisher's choice of n allows the consumers to gain information about the actual content quality through their sample experience before making the purchase decision.

nates the indirect effect. Once $\tilde{V}'(n)$ is sufficiently large, the indirect effect is stronger than the direct effect so that $\frac{\partial D^E}{\partial n} > 0$ and sampling has a *demand-enhancing effect*. In line with Bawa and Shoemaker (2004), we refer to the direct effect of sampling on content demand as the "cannibalization effect" and to the indirect effect as the "expansion effect."

The publisher receives revenues from two sources: selling paid content and including advertisements in the free articles. Specifically, we assume that each of the free articles is supplied with an advertisement. Let $a(n)$ be the inverse advertising demand function, which means that the publisher's choice of n determines the market price for advertisements $a(n)$. Therefore, the publisher's advertising revenues are $a(n)n$. We make the natural assumption that the price for advertisements decreases in sample size, that is, $a'(n) < 0$.

The publisher makes pricing and sampling decisions so as to maximize its (expected) profits from the two sources of revenue:

$$\begin{aligned} \max_{p,n} \quad & \pi(p,n) = (p - c_s)D(p,n,\tilde{V}(n)) + a(n)n \\ \text{s.t.} \quad & p \geq 0 \\ & 0 \leq n \leq N. \end{aligned} \tag{2}$$

Assuming that the publisher's profit function $\pi(p,n)$ is concave and because the constraint set is convex, standard optimization theory posits that there is a unique constraint global maximizer (p^*, n^*). Depending on the optimal pricing and sampling decision, the following definition gives the strategies available to the publisher.

Definition 1 (Strategies). *Given the optimal pricing and sampling decision (p^*, n^*), the publisher adopts either (i) a "sampling strategy" if $p^* > 0$ and $n^* \in (0,N)$, (ii) a pure "paid content strategy" if $p^* > 0$ and $n^* = 0$, or (iii) a pure "free content strategy" if $p^* = 0$ and $n^* = N$.*

Notice that both the paid content strategy and the free content strategy are nested within the sampling strategy: The publisher receives no advertising revenue under a paid content strategy and no sales revenue under a free content strategy. The following result describes the optimal strategy as the ratio of advertising revenue to sales revenue.

Proposition 1 (Advertising-Sales Revenue Ratio). *Under a sampling strategy, the publisher's optimal ratio of advertising revenue to sales revenue is given by*

$$\frac{an^*}{Dp^*} = \frac{\eta_n - \eta_{\tilde{V}}\varepsilon_{\tilde{V}}}{(1 - \frac{1}{\eta_a})\eta_p}, \tag{3}$$

where $\eta_p \equiv -(\partial D/\partial p)(p/D)$ denotes the elasticity of content demand with respect to price, $\eta_n \equiv -(\partial D/\partial n)(n/D)$ denotes the elasticity of content demand with respect to sample size, $\eta_{\tilde{V}} \equiv (\partial D/\partial \tilde{V})(\tilde{V}/D)$ denotes the elasticity of content demand with respect to quality, $\varepsilon_{\tilde{V}} \equiv \tilde{V}'(n)(n/\tilde{V})$ denotes the elasticity of posterior quality expectations with respect to sample size, and $\eta_a \equiv -n'(a)(a/n)$ denotes the price elasticity of advertising demand.

This result has two important managerial insights: First, it shows that the publisher's advertising-sales revenue ratio is determined by characteristics of both the content market and the advertising market. Consumer preferences determine the characteristics of the content market, captured by the elasticities of content demand with respect to price, sample size, and quality. The price elasticity of advertising demand reflects advertiser preferences. This general result thus provides guidance for managers seeking to better understand the contributions of sales and advertising to total revenue.

Second, Proposition 1 shows how changes in the "market environment," captured by the various elasticities, will affect the publisher's composition of revenues. Unsurprisingly, if the price elasticity η_p increases, the advertising-sales revenue ratio is lower. Intuitively, for a given sample size, the optimal price for the content is lower, which results in a higher sales revenue. In contrast, a higher elasticity of content demand with respect to the sample size η_n increases the advertising-sales revenue ratio. Furthermore, the higher the price elasticity of advertising demand η_a, the lower is the advertising-sales revenue ratio.

Proposition 1 also highlights the crucial role which the elasticity of posterior quality expectations with respect to sample size plays. Because the elasticity of content with respect to quality $\eta_{\tilde{V}}$ is positive, the impact of sampling on posterior quality determines the sign of $\eta_{\tilde{V}}\varepsilon_{\tilde{V}}$. Thus, if $\varepsilon_{\tilde{V}}$ is negative, the ratio of advertising revenue to sales revenue tends to be high, while it tends to be low if $\varepsilon_{\tilde{V}}$ is positive. Intuitively, if $\varepsilon_{\tilde{V}} < 0$, sampling reduces expected content demand as $\tilde{V}'(n) < 0$, and hence the advertising-sales revenue ratio is high. In contrast, if $\varepsilon_{\tilde{V}} > 0$, sampling increases expected content demand as consumer revise their expectations about quality upwards, resulting in a lower advertising-sales revenue ratio.

Interestingly, the optimal advertising-sales revenue ratio is reminiscent of the well-known Dorfman-Steiner condition, which states that a monopolist's ratio of advertising spending to sales revenue is equal to the ratio of the elasticities of demand with respect to advertising and price (Dorfman and Steiner, 1954). Proposition 1 reduces to this result in the special case when offering additional samples does not affect posterior quality ($\varepsilon_{\tilde{V}} = 0$) and if the advertising demand is perfectly elastic ($\eta_a \to \infty$).

		Assumptions and Implications		Explicit Form
		Variables	Assumed Properties	
Publisher	Content Parameters	Expected Posterior Quality		
	N . . . content size	$\tilde{V}(n)$	$\tilde{V}'(n) \gtrless 0$	Section 5.2
	$\overline{V}$ . . . maximum content quality	Expected Content Demand		
		$D^E(p,n) \equiv D(p,n,\tilde{V}(n))$	$D_p < 0$, $D_n < 0$, $D_{\tilde{V}} > 0$	Lemma 3
	Cost Parameters		$D^E_n > 0$ . . . demand-enhancing sampling	Lemma 5
	c . . . unit production costs			
	F . . . fixed production costs		$D^E_n < 0$ . . . demand-reducing sampling	Lemma 5
	c_s . . . unit distribution costs			
		Decision Variables		
		p . . . content price		
		n . . . sample size		
Advertiser	Advertiser Parameters	Inverse Advertising Demand		
	A . . . number of advertisers	$a(n)$ . . . free content	$a'(n) < 0$	Section 3.2
	ϕ . . . advertising effectiveness	$a_p(N-n)$ . . . paid content	$a'_p(N-n) < 0$	Section 6.1
Consumer	Prior Parameters	Indirect Utility		Section 3.3
	$\bar{v}_0$ . . . minimum estimate of $\overline{V}$	$u(p,n)$		
	α . . . uncertainty about $\bar{v}_0$	Conditional Indirect Utility		Section 6.2
	Posterior Parameters	$u_i(x)$		
	$\tilde{v}_0(n)$			
	$\alpha + n$			
	Preference Parameters			
	θ . . . valuation of quality			
	ξ . . . ad attraction / ad repulsion			
	x . . . preferred product characteristic			
	τ . . . sensitivity to mismatch			

Table 1: Components of the General Framework

Our general framework is agnostic about how consumers form posterior expectations. To shed light on effects of sampling on posterior quality expectations and in turn expected content demand, the next section introduces a Bayesian learning mechanism in which consumers update their prior expectations about content quality through their sample experience. In order to generate additional insights, we use specific functional forms for content demand and advertising demand. Table 1 summarizes the main model assumptions (as well as its core components) and indicates where the reduced-form expressions are derived analytically.

3 Model

This section introduces the components of our model. We begin by laying out the assumptions regarding the publisher and the advertisers. We then describe how consumers learn about content quality. Finally we lay out the timeline of the model.

3.1 The Publisher

The publisher offers an information good with $N \in \mathbb{N}$ content parts whose qualities are uniformly distributed on the quality spectrum $[0, \overline{V}]$, where $\overline{V}$ is the publisher's private information. The publisher allows the consumers to sample n out of the N content parts ($n \leq N$). The qualities of the n free samples are also uniformly distributed on $[0, \overline{V}]$ and are labeled $V_1, \ldots, V_n$. We normalize both the marginal costs c of producing the content and the costs of providing digital access c_s to zero.

3.2 Advertisers

There are A advertisers who differ in the willingness to pay for a placing their ads in the free articles offered by the publisher. Such heterogeneity might reflect differences in profits from selling their advertised products. To capture this heterogeneity, we assume that advertisers' willingness to pay $\hat{\phi}$ is drawn independently from a uniform distribution over the interval $[0, \phi]$. In this setting, advertising demand as a function of ad price a can be derived as

$$n(a) = A \Pr\left\{\hat{\phi} \geq a\right\} = A\left(1 - \frac{a}{\phi}\right).$$

The inverse advertising demand, which maps the publisher's choice of n to the market price for advertisements $a(n)$, is thus given by

$$a(n) = \phi\left(1 - \frac{n}{A}\right).$$

This function slopes downward, implying that the price the publisher receives for an ad is decreasing in the number of free articles offered. Further, the inverse demand exhibits the natural properties that advertising prices are increasing in both the market size A and the maximum willingness to pay to place an advertisement ϕ.

To obtain a parsimonious specification of demand, we impose the normalization $\phi = \frac{A}{N}$, which allows to write the inverse advertising demand as

$$a(n) = \phi - \frac{n}{N}, \tag{4}$$

where we assume that $\phi > 1$. Adopting the terminology of Godes et al. (2009), we will refer to ϕ as "advertising effectiveness." Basically, ϕ can be thought of as a parameter shifting the (inverse) demand function "outwards."

3.3 Consumers

Consumers know that the qualities of the free samples are uniformly distributed on the interval $[0, \overline{V}]$, but they do not know the upper bound of the publisher's quality spectrum $\overline{V}$ and are hence uncertain about (average) content quality.[5] Consumers do have a common prior belief about $\overline{V}$ that may stem, for instance, from reviews, ratings or "word of mouth." The natural conjugate family for a random sample from a uniform distribution with unknown upper bound is the Pareto distribution (DeGroot, 1970). We capture uncertainty about $\overline{V}$ by a prior belief that consists of a minimum estimate $\bar{v}_0$ of the upper bound $\overline{V}$ and a level of uncertainty α about this value. Specifically, we assume that the prior belief follows a Pareto distribution with density function

$$f(\bar{v}|\bar{v}_0, \alpha) = \begin{cases} \dfrac{\alpha \bar{v}_0^{\alpha}}{\bar{v}^{\alpha+1}}, & \text{for } \bar{v} > \bar{v}_0 \\ 0, & \text{otherwise.} \end{cases}$$

We assume that $\alpha > 1$ to ensure existence of the prior expectations.[6] Further, we assume that the consumers' prior parameters $\bar{v}_0$ and α are common knowledge. For instance, the publisher can learn about prior expectations by employing standard market research techniques such as surveys. Based on the consumers' prior knowledge about $\bar{v}_0$ and α, their prior expectation about $\overline{V}$ is

$$\mathbb{E}[\overline{V}|\bar{v}_0, \alpha] = \frac{\alpha \bar{v}_0}{\alpha - 1}. \tag{5}$$

[5] Note that the upper bound $\overline{V}$ is monotonically related to the mean, which may be an alternative way for consumers to think about content quality.

[6] Our measure of uncertainty corresponds to the scale parameter α of the Pareto distribution. Hence, when the uncertainty is higher, the prior distribution is more spread out.

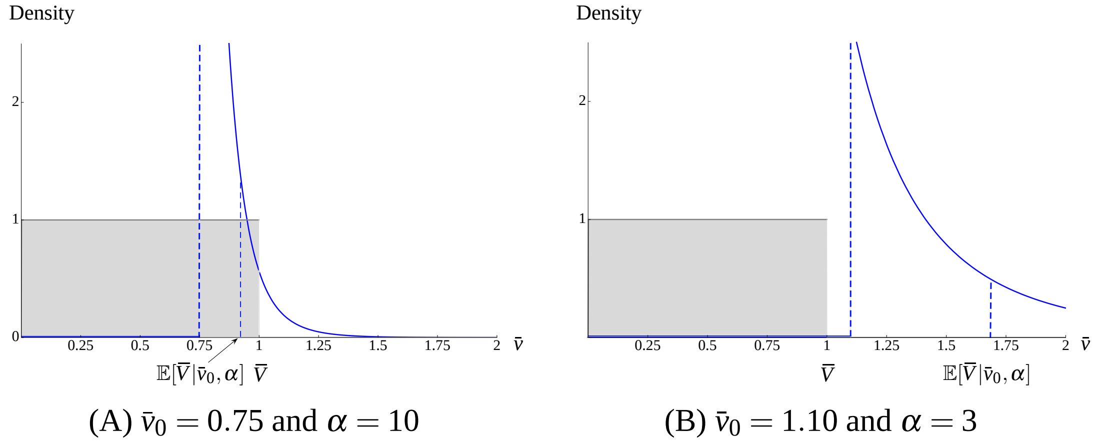


Figure 1: Prior expectations about $\overline{V}$ (where $\overline{V} \equiv 1$).

Obviously, prior expectations increase in $\bar{v}_0$ and decrease in α. Figure 1 illustrates prior beliefs along with the corresponding expectations for different parameter values. Prior expectations are lower than actual quality in Panel A and higher than actual quality in Panel B. Note that prior expectations can be higher than actual quality even if $\bar{v}_0 < \overline{V}$.

Consumers update their prior belief about $\overline{V}$ by taking the observed qualities of the free samples into account. Specifically, consumers evaluate the n sample qualities $V_i = v_i$ ($i = 1,\ldots,n$) to form their posterior beliefs $\tilde{v}(n)$ about $\overline{V}$. Using standard Bayesian analysis, $\tilde{v}(n)$ follows a Pareto distribution with minimum value parameter $\tilde{v}_0(n) = \max\{\bar{v}_0, v_1, \ldots, v_n\}$ and shape parameter $\alpha + n$ (De Groot, 1970).[7] Hence, the posterior expectation of $\overline{V}$ is given by

$$\mathbb{E}[\overline{V}|\tilde{v}_0(n),\alpha] = \frac{(\alpha+n)\tilde{v}_0(n)}{\alpha+n-1}.$$

Consumers infer the expected quality of the information good $\mathbb{E}[V|v_1,\ldots,v_n]$ from the average quality of the sampled content parts. Knowing that qualities are uniformly distributed on the quality spectrum offered, the expected quality of the information good is given by

$$\mathbb{E}[V|v_1,\ldots,v_n] = \frac{\mathbb{E}[\overline{V}|\tilde{v}_0(n),\alpha]}{2}. \tag{6}$$

Consumers agree that higher quality is better than lower quality but differ in the way they value quality. To capture this heterogeneity, we introduce a preference parameter for quality θ, which is uniformly distributed on the interval [0,1]. We consider discrete choice and assume that each consumer either purchases the information good at price p

[7] The proof of this result is reproduced in the Appendix.

or stays with the n free samples. A consumer's indirect utility from these two options is given by

$$u(p,n) = \begin{cases} \theta N \mathbb{E}[V|v_1,\ldots,v_n] + \xi n - p, & \text{from purchasing at price } p \\ \theta n \mathbb{E}[V|v_1,\ldots,v_n] + \xi n, & \text{from staying with the free samples,} \end{cases}$$

where ξ denotes a consumer's respective intensity of ad-attraction ($\xi > 0$) or ad-repulsion ($\xi < 0$). Thus, when a consumer exhibits ad-loving behavior, the utility of both options is augmented by ξn, while the utility of both options is reduced by ξn in the case of ad-avoiding behavior (see, for instance, Gabsezwicz et al. 2004).

In this utility function the value of the information good is equal to the number of content parts multiplied by the expected quality.[8] This implies that a consumer will purchase the information good if and only if the indirect utility from buying exceeds the indirect utility from consuming the free samples only, that is, if

$$\theta(N-n)\mathbb{E}[V|v_1,\ldots,v_n] - p \geq 0. \tag{7}$$

This condition means that the value of the content that has not been sampled must exceed the price. Importantly, the purchase condition does not depend on consumer behavior towards advertising.

3.4 Timeline

The publisher first decides on the sample size n and the price p at which to sell the information good. Next, consumers select the samples of their choice and use the observed sample qualities $V_1 = v_1, \ldots, V_n = v_n$ to update their prior expectations about content quality $\overline{V}$. Finally, consumers decide whether or not to purchase the information good based on posterior expectations.

4 Strategy with Known Quality

We first analyze as a benchmark the case in which the consumers know $\overline{V}$ and hence the publisher's quality spectrum. In this setting, sampling does not affect the consumers' expectations about quality and simply serves to generating advertising revenues. We derive content demand for each strategy and then characterize the optimal strategy.

[8]This additivity assumption is justified for independently valued content parts. However, a concave or convex relationship between the value and the number of content parts might be more appropriate for interrelated content parts, that is, if the content parts are substitutes or complements.

4.1 Content Demand

We first derive content demand under a sampling strategy and subsequently the demands for the two boundary strategies.

Sampling Strategy. When consumers know the upper bound $\bar{V}$ of the quality spectrum, they expect content quality to be equal to $\mathbb{E}[V] = \frac{\bar{V}}{2}$. When the publisher employs the sampling strategy, consumers get some content for free but have to purchase the information good if they want to obtain the full content. Because θ follows a uniform distribution on $[0,1]$, the purchase condition in (7) implies that the content demand can be expressed as

$$\begin{aligned} D(p,n) &= \Pr\left\{\theta \geq \frac{p}{(N-n)\frac{\bar{V}}{2}}\right\} \\ &= \max\left\{0, 1 - \frac{p}{(N-n)\frac{\bar{V}}{2}}\right\}. \end{aligned} \tag{8}$$

This demand function has the intuitive properties that it decreases in price p and increases in average quality $\frac{\bar{V}}{2}$. Moreover, sampling has a direct negative effect on content demand as $\frac{\partial D}{\partial n} < 0$. Intuitively, this follows because a larger sample size reduces the utility of the remaining content consumers have to pay for.

Paid Content Strategy. When the publisher employs a paid content strategy, setting $n = 0$ in (8) produces

$$D(p,0) = \max\left\{0, 1 - \frac{p}{N\frac{\bar{V}}{2}}\right\}. \tag{9}$$

Free Content Strategy. When the publisher employs a free content strategy, consumers do not purchase the information good as they can download it for free and hence $D(p,N) \equiv 0$.

4.2 Optimal Pricing and Sampling

The publisher's makes its pricing and sampling decision so as to

$$\begin{aligned} \max_{p,n} \quad & \pi(p,n) = p\left(1 - \frac{p}{(N-n)\frac{\bar{V}}{2}}\right) + \left(\phi - \frac{n}{N}\right)n \\ \text{s.t.} \quad & p \geq 0 \\ & 0 \leq n \leq N. \end{aligned}$$

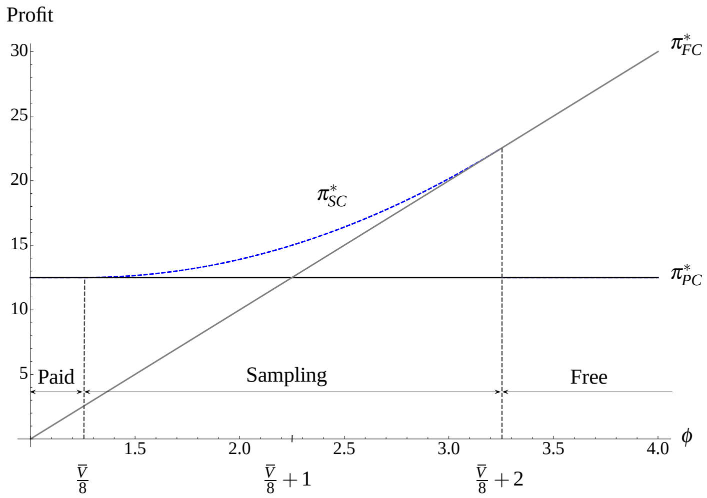


Figure 2: Optimal strategy with known quality (for $\bar{V} = 10$ and $N = 10$).

From the first-order conditions, the optimal price for a given sample size is

$$p(n) = \frac{(N-n)\bar{V}}{4}. \tag{10}$$

This implies that the more free samples the publisher chooses to offer, the less he will be able to charge the consumer for the remaining content. The next result summarizes the optimal pricing and sampling decisions for each of the three strategies.

Lemma 1 (**Pricing and Sampling**). *Suppose that the upper bound of content quality $\bar{V}$ is common knowledge. Then, (i) under a sampling strategy, $p^* = N\bar{V}(8(2-\phi)+\bar{V})/64$ and $n^* = N(8\phi - \bar{V})/16$, (ii) under a paid content strategy, $p^* = N\bar{V}/4$ and $n^* = 0$, and (iii) under free content strategy, $p^* = 0$ and $n^* = N$.*

The parameters $\bar{V}$ and ϕ have opposite effects on the optimal price and on the optimal sample size under a sampling strategy: As we can expect, p^* increases in $\bar{V}$ while n^* decreases in the highest quality. In contrast, p^* decreases in ϕ, and n^* increases in advertising effectiveness. Both the optimal price and the optimal sample size increase in content size N.

The following proposition characterizes the publisher's optimal strategy.

Proposition 2 (**Optimal Strategy**). *If consumers know the quality spectrum, then: (i) if $\phi \in (\frac{\bar{V}}{8}, \frac{\bar{V}}{8}+2)$, the publisher should employ a sampling strategy, (ii) if $\phi \leq \frac{\bar{V}}{8}$, the publisher should follow a paid content strategy, and (iii) if $\phi \geq \frac{\bar{V}}{8}+2$, the publisher's optimal strategy is a free content strategy.*

Proposition 2 shows that the choice of the optimal strategy is driven by the relationship between content quality $\bar{V}$ and advertising effectiveness ϕ. Thus, for a given content quality, a paid content strategy is optimal if the effectiveness of advertising is sufficiently low. For intermediate levels of advertising effectiveness, a sampling strategy that generates revenues from both sales and advertising on the free samples is optimal. If advertising is sufficiently effective, the publisher should switch to a free content strategy. Figure 2 illustrates the optimal strategy for different values of ϕ and $\overline{V}$ along with the profits for each strategy (π^*_{SC} for the sampling strategy, π^*_{PC} for the paid content strategy, and π^*_{FC} for the free content strategy).[9]

The effects of ϕ and $\overline{V}$ on the optimal strategy can also be understood by inspection of the advertising-sales revenue ratio. The ratio follows from (3) and is

$$\frac{an^*}{Dp^*} = \frac{(\phi - \frac{\overline{V}}{8})(\frac{\overline{V}}{8} + \phi)}{\frac{\overline{V}}{4}(\frac{\overline{V}}{8} + 2 - \phi)}.$$

The ratio of advertising revenue to sales revenue tends to zero as ϕ approaches the lower bound $\frac{\overline{V}}{8}$, implying that the publisher should employ a paid content strategy. A sampling strategy is optimal only if advertising is not "too effective," that is, as long as $\phi \leq \frac{\overline{V}}{8} + 2$. Once ϕ exceed this level, the publisher should switch to a free content strategy.

4.3 Summary

When content quality is common knowledge, the publisher's optimal strategy is solely determined by the relation between advertising effectiveness and content quality. The more effective advertising is, the more free samples the publisher should offer—even though it cannibalizes content demand. In the next section, we study optimal pricing and sampling decisions when the quality spectrum is not known to consumers who learn about quality through inspection of free samples.

5 Strategy with Unknown Quality

When $\overline{V}$ and hence the product spectrum is not known to consumers, sampling not only serves the purpose of generating advertising revenues but also influences consumers' expectations about quality. As in the benchmark model, we first derive content demand for each strategy and then characterize the optimal sampling strategy.

5.1 Content Demand

We first derive content demand under a sampling strategy and subsequently derive the demands for the two boundary strategies.

[9] Qualitatively, the choice of specific parameter values does not affect Figure 2.

Sampling Strategy. When consumers do not know the upper bound of the quality spectrum $\overline{V}$ with certainty, content demand is influenced by consumers' posterior quality expectations. Therefore, when the publisher makes decisions about the sample size and the price, it has to base them on *expected* content demand as consumers have not yet evaluated sample qualities and updated their expectations about content quality.

Calculating expected content demand involves a two-step procedure. In the first step, the publisher computes the expected posterior quality by averaging posterior expectations about $\overline{V}$ as given in (6) across *all* possible realizations of sample qualities:

$$\mathbb{E}\left[\mathbb{E}[V|V_1,\ldots,V_n]\right] = \frac{(\alpha+n)\,\mathbb{E}\left[\tilde{v}_0(n)\right]}{2\,(\alpha+n-1)}.$$

In the second step, the publisher substitutes the expected posterior quality into the purchase condition given in (7) to obtain expected content demand:

$$D^E(p,n) = \max\left\{0, 1-\frac{p}{(N-n)}\frac{2\,(\alpha+n-1)}{(\alpha+n)\,\mathbb{E}\left[\tilde{v}_0(n)\right]}\right\}. \tag{11}$$

Next, we calculate $\mathbb{E}\left[\tilde{v}_0(n)\right]$ and insert it into the expected content demand given in (11). The following lemma summarizes the result.

Lemma 2 (**Expected Demand**)**.** *When the publisher sells the information good at price p and offers $n \in \{1, N-1\}$ samples, then*

(a) if $\bar{v}_0 < \overline{V}$, expected content demand is given by

$$D^E_{\{\bar{v}_0<\overline{V}\}}(p,n) = \max\left\{0, 1-\frac{p}{(N-n)}\frac{2\,(\alpha+n-1)\,(n+1)\overline{V}^n}{(\alpha+n)\left(\bar{v}_0^{n+1}+n\overline{V}^{n+1}\right)}\right\}. \tag{12}$$

(b) if $\bar{v}_0 \geq \overline{V}$, expected content demand is given by

$$D^E_{\{\bar{v}_0\geq\overline{V}\}}(p,n) = \max\left\{0, 1-\frac{p}{(N-n)}\frac{2\,(\alpha+n-1)}{(\alpha+n)\,\bar{v}_0}\right\}. \tag{13}$$

These demand functions have the intuitive properties that they decrease in price p and increase in expected posterior quality. Hence, sampling has both a direct demand-reducing effect and an indirect effect that operates through its impact on posterior expectations. The direct effect kicks in through the factor $\frac{1}{N-n}$ and mirrors the cannibalization effect $\frac{\partial D}{\partial n} < 0$.

Paid Content Strategy. When the publisher employs a paid content strategy, consumers cannot update their quality expectations. Setting $n = 0$ in (11) and rearranging produces

$$D^E(p,0) = \max\left\{0, 1-\frac{p}{N\frac{\alpha\bar{v}_0}{2(\alpha-1)}}\right\}. \tag{14}$$

This demand function is a close cousin of the demand for paid content in (9) when consumers know quality. The difference is that the expected content demand is driven by prior expectations about $\overline{V}$ rather than expected quality $\frac{\overline{V}}{2}$ itself.

Free Content Strategy. When the publisher employs a free content strategy, consumers do not purchase the information good as they can download it for free and hence $D^E(p,N) \equiv 0$.

5.2 The Role of Quality Expectations

For a given level of prior expectations about content quality, sampling either increases or decreases expected content demand. Whether or not sampling compensates for cannibalization through consumers' learning depends on the gap between posterior quality and actual quality. We define expected posterior quality as

$$\tilde{V}(n) = \begin{cases} \dfrac{(\alpha+n)\left(\bar{v}_0^{n+1}+n\overline{V}^{n+1}\right)}{2(\alpha+n-1)(n+1)\overline{V}^n}, & \text{if } \bar{v}_0 < \overline{V} \\ \dfrac{(\alpha+n)\bar{v}_0}{2(\alpha+n-1)}, & \text{if } \bar{v}_0 \geq \overline{V} \end{cases} \tag{15}$$

and the quality gap as $\tilde{V}(n) - \frac{\overline{V}}{2}$. Consumers overestimate (underestimate) quality if the expected posterior quality is higher (lower) than actual quality. This leads to the following result.

Lemma 3 (Quality Gap). *When $\bar{v}_0 < \overline{V}$, consumers overestimate quality after their sample experience if*

$$\frac{\bar{v}_0}{\overline{V}} > \left(\frac{\alpha-1}{\alpha+n}\right)^{\frac{1}{n+1}}, \tag{16}$$

and underestimate it if the inequality is reversed. If $\bar{v}_0 \geq \overline{V}$, consumers overestimate quality irrespective of the sample size and their level of uncertainty $\alpha > 1$.

The intuition behind this result is perhaps best understood by recalling that prior expectations can be higher than actual quality even if $\bar{v}_0 < \overline{V}$ (a high level of uncertainty about $\bar{v}_0$ is captured by a low α). Condition (16) applies when consumers overestimate quality based on posterior expectations: This is likely to be the case for a low α and when the publisher offers a small number of free articles n. On the other hand, consumers underestimate quality if their uncertainty is low and the sample size is large.

For the case where $\bar{v}_0 < \overline{V}$, Figure 3 illustrates the set of prior parameters for which consumers overestimate and underestimate quality, respectively. The latter parameter

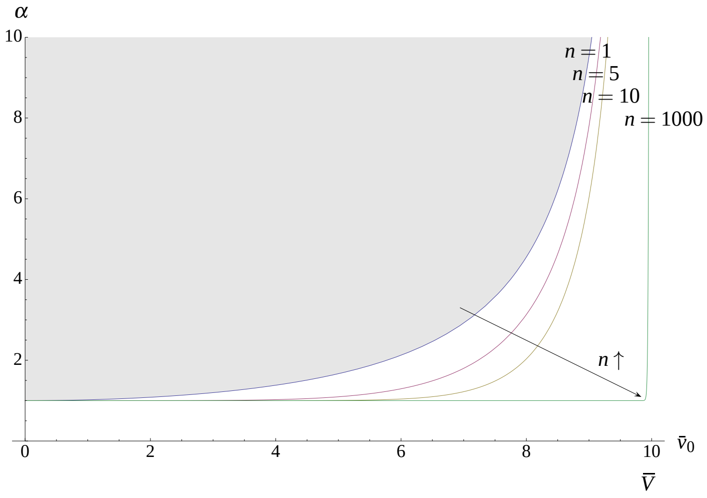


Figure 3: The quality gap for the case $v_0 < \overline{V}$ (where $\overline{V} \equiv 10$). The shaded area indicates where consumers underestimate quality.

region $(\bar{v}_0, \alpha)$ is indicated by the shaded area. By construction, where $\alpha > 1$, condition (16) holds and consumers overestimate quality. The figure also illustrates that the parameter region for which consumers overestimate quality shrinks as n gets larger. Formally, this can be seen by noting that $\tilde{V}(n) \rightarrow \frac{\overline{V}}{2}$ as $n \rightarrow \infty$, meaning that consumers learn actual quality once the sample size gets "large enough."

The definition of $\tilde{V}(n)$ allows us to rewrite the expected content demand derived in Lemma 2 more compactly as

$$D^E(p,n) = \max\left\{0, 1 - \frac{p}{(N-n)\tilde{V}(n)}\right\}. \tag{17}$$

Notice that this is a specific form of the reduced-form demand function in Equation (1). Hence the number of free samples n has both a direct effect on expected content demand and an indirect effect that operates through posterior quality expectations $\tilde{V}(n)$. The next result uses this demand function to identify conditions under which sampling has a demand-enhancing effect (that is, $\frac{\partial D^E}{\partial n} > 0$).

Lemma 4 (**Effects of Sampling**). *Offering free samples has a demand-enhancing effect if $\varepsilon_{\tilde{V}} > \frac{n}{N-n}$, that is, if the elasticity of consumers' posterior quality expectations exceeds the ratio of sampled to paid content.*

Lemma 4 shows that offering free samples may increase expected content demand through consumers' learning, even though it produces a cannibalization effect. Intuitively, the indirect effect dominates the direct cannibalization effect if sampling induces a sufficiently large upwards revision of consumers' prior expectations.

5.3 Optimal Strategy

The publisher's makes its pricing and sampling decisions so as to

$$\begin{aligned}\max_{p,n} \quad & \pi^E(p,n) = p\left(1 - \frac{p}{(N-n)\tilde{V}(n)}\right) + \left(\phi - \frac{n}{N}\right) n \\ \text{s.t.} \quad & p \geq 0 \\ & 0 \leq n \leq N.\end{aligned}$$

The only difference between this expected profit and the profits when content quality is known to consumers is the dependence on expected posterior quality rather than actual (average) quality. Based on a comparison to (10) and recalling that $\tilde{V}(n)$ is the posterior estimate of average quality $\frac{\overline{V}}{2}$, we thus obtain that

$$p(n) = \frac{(N-n)\tilde{V}(n)}{2}.$$

Substituting $p(n)$ back into the profit function allows us to rewrite the profit maximization problem as

$$\begin{aligned}\max_{n} \quad & \pi^E(n) = (N-n)\frac{\tilde{V}(n)}{4} + \left(\phi - \frac{n}{N}\right) n \\ \text{s.t.} \quad & 0 \leq n \leq N.\end{aligned} \tag{18}$$

In contrast to our benchmark model, it is not possible to characterize the optimal pricing and sampling decisions (and hence profits) analytically. Nevertheless, we have the following result.

Proposition 3 (Optimal Strategy). *Suppose that consumers are uncertain about $\overline{V}$ and that the profit function $\pi^E(n)$ is strictly concave. Then, there are cut-off values $\underline{\phi} = \frac{1}{4}(\tilde{V}(0) - N\tilde{V}'(0))$ and $\overline{\phi} = 2 + \frac{\tilde{V}(N)}{4}$ such that a sampling strategy is optimal for $\phi \in (\underline{\phi}, \overline{\phi})$, a paid content strategy is optimal for $\phi \leq \underline{\phi}$, and a free content strategy is optimal for $\phi \geq \overline{\phi}$.*

This result is consistent with the insights from the benchmark model: a paid content strategy is optimal only if the advertising effectiveness is sufficiently low, a sampling strategy is optimal for intermediate levels of the advertising effectiveness, and the publisher should switch to a free content strategy once advertising is sufficiently effective (see Proposition 2).[10] Figure 4 illustrates the optimal strategy for varying advertising effectiveness ϕ and the expected profits for each strategy (π^E_{SC} for the sampling strategy, π^*_{PC} for the paid content strategy, and π^*_{FC} for the free content strategy).

[10]Observe that we assume in Proposition 3 that the profit function $\pi^E(n)$ is globally concave. However, there are parameter constellations for which this assumption is not satisfied. In this case, the cut-off values must be determined numerically by the comparing profits that arise from the different strategies.

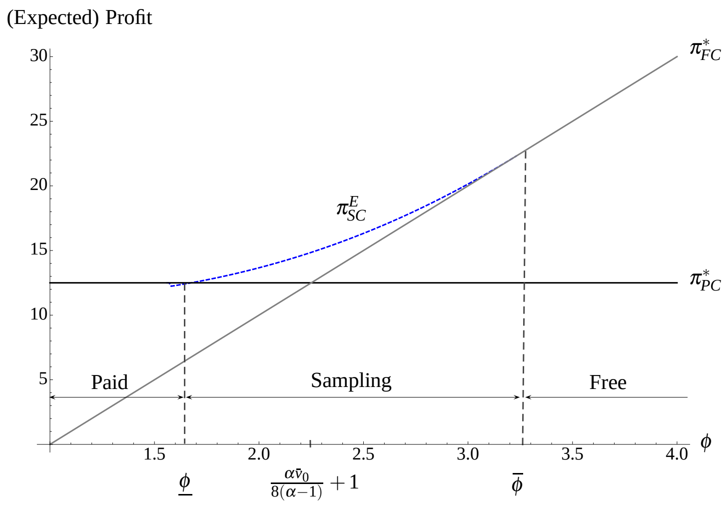
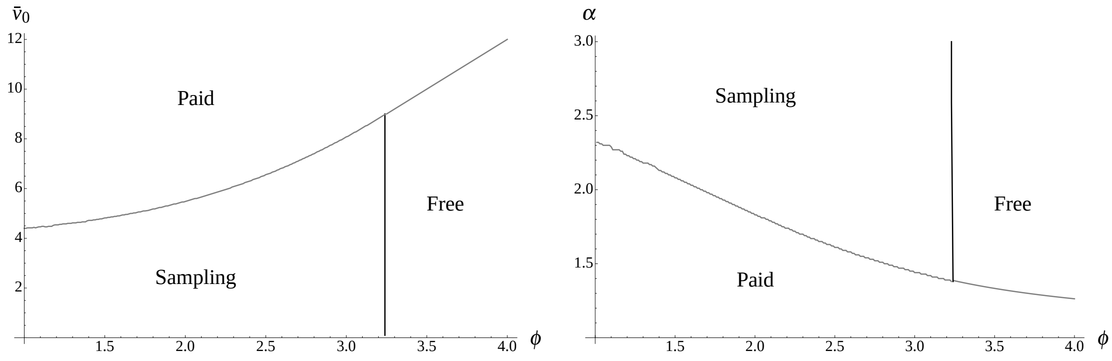


Figure 4: Optimal strategy with unknown quality (for $\bar{v}_0 = 5$, $\alpha = 2$, $\overline{V} = 10$, and $N = 10$).

Proposition 3 reveals that prior expectations determine the lower of the two cut-off values for a sampling strategy to be optimal whereas posterior expectations for sample size $n = N$ determine the upper cut-off value. In effect, $\underline{\phi}$ is determined by the impact of the "first" free content part on posterior expectations, while $\overline{\phi}$ is determined by posterior expectations after inspection of the "last" free content part. The next lemma shows that the model where quality $\overline{V}$ is not known to consumers nests the full information benchmark model (see Proposition 2).

Lemma 5 (**Cut-off Values**)**.** *Suppose that consumers are uncertain about content quality $\overline{V}$ and that the profit function $\pi^E(n)$ is strictly concave. Then, when consumers have correct quality expectations, that is, if $\bar{v}_0 = \overline{V}$ and $\alpha \to \infty$, the lower bound $\underline{\phi}$ converges to $\frac{\overline{V}}{8}$ and the upper bound $\overline{\phi}$ converges to $\frac{\overline{V}}{8} + 2$.*

We next explore the comparative statics effect of changes in the consumer's prior parameters on the optimal strategy. Proposition 3 shows that the optimal strategy depends not only on advertising effectiveness ϕ and quality $\overline{V}$ as in the benchmark model, but also on the specific values of the prior parameters $\bar{v}_0$ and α (as well as content size N). Figure 5 illustrates the effects of changes in the consumers' prior parameters. Panel A depicts the cut-off thresholds between the different strategies in the $(\bar{v}_0, \phi)$-space (given $\alpha = 2$). Similarly, Panel B illustrates the optimal choice of strategy in the (α, ϕ)-space (given $\bar{v}_0 = 5$). Here prior expectations are correct and coincide with actual quality when $\bar{v}_0 = 5$ and $\alpha = 2$ (see Equation 5). The following observation summarizes our insights.



(A) Based on $\bar{v}_0$ and ϕ (B) Based on α and ϕ

Figure 5: Optimal strategy (for $\overline{V} = 10$ and $N = 10$).

Observation 1 (**Comparative Statics**). *Suppose that consumers are uncertain about quality $\overline{V}$. Then, (a) when both prior quality expectations and advertising effectiveness are low, the publisher should employ a sampling strategy to reveal his higher than expected quality, (b) when prior expectations increase, that is, either $\bar{v}_0$ increases or α decreases, the publisher should switch to a paid content strategy, and (c) when the advertising effectiveness ϕ increases sufficiently, the publisher should adopt a free content strategy.*

5.4 Summary

When content quality is the publisher's private information, sampling has a demand-enhancing effect when the elasticity of consumer's posterior expectations with respect to sample size exceeds the ratio of sampled to paid content. When this condition is not satisfied, sampling mitigates or reinforces the cannibalization effect. As in the benchmark model, we show that employing a paid content strategy is optimal only if advertising effectiveness is sufficiently low compared to prior quality expectations, a sampling strategy is optimal for intermediate levels of advertising effectiveness, and the publisher should switch to a free content strategy once advertising is sufficiently effective compared to posterior quality expectations.

6 Model Extensions

This section extends our model in two ways. The first extension allows for the inclusion of advertisements in both the free articles and the paid content, while the second extension introduces competition among publishers.

6.1 Including Advertisements in the Paid Content

In this section, we extend the model by allowing it to include advertisements in both the free articles and the paid content. To this end, we assume that the market price for advertisements included in the paid content is given by

$$a_p(\hat{n}) = \phi_p - \frac{\hat{n}}{N},$$

where $\hat{n} \equiv N - n$ and $\phi_p > 1$ denotes the advertising effectiveness for ads in the paid content. This inverse demand is a natural counterpart to the advertising demand $a(n)$ given in (4) and reflects that the ad price depends number of articles $\hat{n}$ that have not been offered as free samples. Differences in the levels of advertising effectiveness ϕ_p and ϕ capture differences in reach or the degree of targeting in the advertising markets for paid and free content.

Allowing for advertisements in the paid content affects content demand. Specifically, content demand now depends on consumer behavior towards advertising. Letting ξ denote a consumer's intensity of ad-repulsion or ad-attraction, the expected content demand can be derived as

$$\hat{D}^E(p,n) = \max\left\{0, 1 - \frac{1}{\mathbb{E}[V|v_1,\ldots,v_n]}\left(\frac{p}{N-n} - \xi\right)\right\}.$$

Compared to the case where consumers are ad-neutral, content demand is higher when consumers are ad-lovers ($\xi > 0$) and lower when consumers are ad-avoiders ($\xi < 0$).

The publisher makes its pricing and sampling decisions so as to

$$\begin{aligned} \max_{p,n} \quad & \pi^E(p,n) = (p + R_p(\hat{n}))\hat{D}^E(p,n) + \left(\phi - \frac{n}{N}\right)n \\ \text{s.t.} \quad & p \geq 0 \\ & 0 \leq n \leq N, \end{aligned}$$

where $R_p(\hat{n}) \equiv a_p(\hat{n})\hat{n}$ are the additional revenues from including ads in the paid content. By definition, $R_p = 0$ under a free content strategy, while $R_p = (\phi_p - 1)N$ under a paid content strategy.

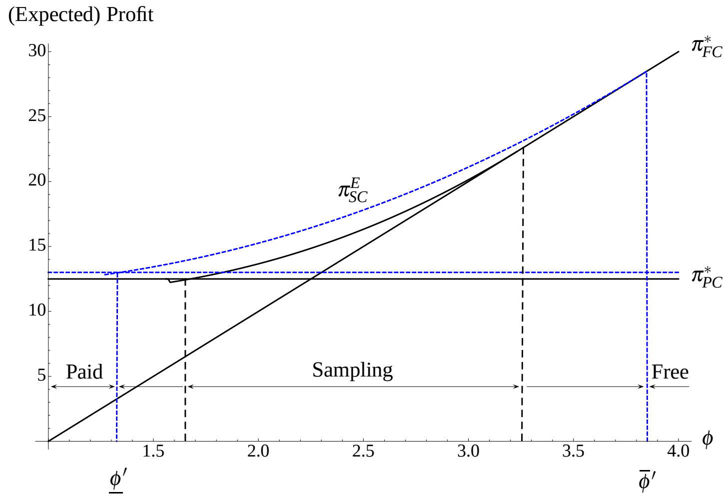


Figure 6: Optimal strategy with (dashed lines) and without advertisements in paid content (solid lines) for $\xi = 0$, $\bar{v}_0 = 5$, $\alpha = 2$, $\overline{V} = 10$, $N = 10$, and $\phi_p = 1.1$.

Figure 6 illustrates the profit effects of including advertisements in the paid content when consumers are neutral about advertisements and shows that the range of advertising effectiveness ϕ for which the sampling strategy is best expands. Intuitively, exploiting revenues from advertisements in the paid content increases the unit margin from selling content, which translates into higher profits under a sampling strategy. These profit effects are more pronounced when advertising effectiveness for ads in paid content ϕ_p increases. Further, the profits under a sampling strategy are higher when the consumers are ad-lovers (for given ϕ_p) and lower when they are ad-avoiders.

6.2 The Impact of Competition

Thus far, we have examined a publisher operating in a monopoly setting. In this section, we allow for competition between two publishers indexed by $i = 1,2$. Both firms choose their sample size n_i and sell their content at price p_i. Horizontal differentiation is à la Hotelling, and we assume that the firms are located at the extremes of the product spectrum at $x_1 = 0$ and $x_2 = 1$, respectively. Vertical differentiation captures the firms' different content qualities.

We again assume that the perceived quality q_i of information good i is equal to its number of content parts N_i multiplied by its expected posterior quality, that is $q_i = N_i \mathbb{E}[V_i | v_{1i}, \ldots, v_{ni}]$. A consumer's indirect utility from buying information good i is given by

$$u_i(x) = q_i - \tau\, |x - x_i| - p_i,$$

		Firm 2		
		PC	*SC*	*FC*
Firm 1	*PC*	π_1^{PP}, π_2^{PP}	π_1^{PS}, π_2^{PS}	π_1^{PF}, π_2^{PF}
	SC	π_1^{SP}, π_2^{SP}	π_1^{SS}, π_2^{SS}	π_1^{SF}, π_2^{SF}
	FC	π_1^{FP}, π_2^{FP}	π_1^{FS}, π_2^{FS}	π_1^{FF}, π_2^{FF}

Figure 7: Strategy choices and corresponding profits.

where $x \in [0,1]$ is the consumer's most preferred product characteristic (drawn independently across consumers from a uniform distribution over the interval $[0,1]$) and the parameter $\tau > 0$ measures the consumer's sensitivity to horizontal mismatch $|x - x_i|$. The location of the indifferent consumer $\hat{x}$ follows from solving the indifference condition $u_1(\hat{x}) = u_2(\hat{x})$ for given prices $\mathbf{p} = (p_1, p_2)$ and sample sizes $\mathbf{n} = (n_1, n_2)$.[11] Content demands are given by

$$D_1(\mathbf{p},\mathbf{n}) = \frac{N_1 - n_1}{N_1}\hat{x}(\mathbf{p},\mathbf{n}) \quad \text{and} \quad D_2(\mathbf{p},\mathbf{n}) = \frac{N_2 - n_2}{N_2}(1 - \hat{x}(\mathbf{p},\mathbf{n})),$$

where $\frac{N_i - n_i}{N_i}$ is the conditional purchase probability given sample size n_i. Consumers thus choose their preferred publisher based on prices and posterior quality expectations and subsequently purchase the content with probability $\frac{N_i - n_i}{N_i}$. The sampling decision n_i therefore has a direct effect on content demand through the conditional purchase probability (a cannibalization effect) and an indirect effect on $\hat{x}(\mathbf{p},\mathbf{n})$ (an expansion effect). Note that publisher i's content demand is zero under a free content strategy due to the cannibalization effect.

Publisher i makes its pricing and sampling decisions so as to

$$\begin{aligned} \max_{p_i, n_i} \quad & \pi_i^E(\mathbf{p},\mathbf{n}) = p_i D_i(\mathbf{p},\mathbf{n}) + \left(\phi_i - \frac{n_i}{N_i}\right) n_i \\ \text{s.t.} \quad & p_i \geq 0 \\ & 0 \leq n_i \leq N_i, \end{aligned}$$

where ϕ_i is the advertising effectiveness of publisher i's advertising. Compared to the monopoly case, each publisher now has to take into account the rival's choice of strategy

[11] See Anderson et al. (1992) for a comprehensive treatment of discrete choice models.

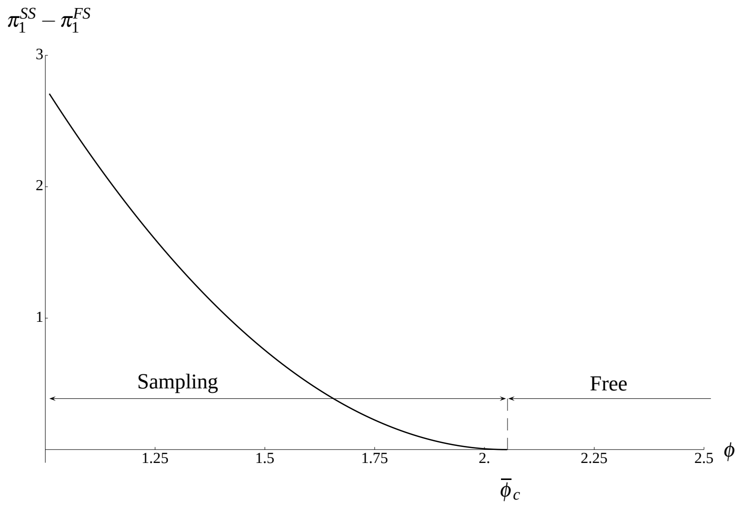


Figure 8: Optimal strategy in the symmetric equilibrium (for $\bar{v}_0 = 0.6$, $\alpha = 2$, $\overline{V}_i = 2$, $N_i = 10$, and $\tau = 1$).

to make its optimal decision. Thus, there are nine possible outcomes in the first-stage game, summarized in Figure 7. If both firms use a paid content strategy, the firms' corresponding profits are denoted by π_1^{PP} and π_2^{PP}, respectively (and likewise for the other outcomes). For each outcome, the profit levels can be obtained by solving the publishers' decision problems. The optimal strategy choice is then obtained as a Nash equilibrium of the first-stage game.[12]

The first-stage game is complex so that little analytical headway can be made. To illustrate the optimal strategy choice, we focus on a market environment in which the paid content strategies are strictly dominated and the publishers can only choose between the two strategies *SC* and *FC*. For example, the strategy combination (*SC*, *SC*) is a Nash equilibrium if neither publisher has an incentive to unilaterally deviate from the sampling strategy. This is the case if the no-deviation constraints $\pi_1^{SS} - \pi_1^{FS}$ (firm 1) and $\pi_2^{SS} - \pi_2^{SF}$ (firm 2) hold.

Figure 8 illustrates the no-deviation condition in a symmetric equilibrium.[13] A sampling strategy is optimal for both publishers if the advertising effectiveness $\phi_i \equiv \phi$ is lower than $\overline{\phi}_c$ and a free content strategy is optimal if $\phi \geq \overline{\phi}_c$. This finding is consistent with the insights from the monopoly setting. Of course, in a setting with competition, there are asymmetric industry configurations in which one publisher employs a free content strategy while the rival uses a sampling strategy.

[12]See Fudenberg and Tirole (1991) for the game theoretic concepts.

[13]Due to symmetry, the no-deviation conditions $\pi_1^{SS} - \pi_1^{FS}$ and $\pi_2^{SS} - \pi_2^{SF}$ are the same.

7 Summary and Implications

This paper analyzed digital content strategies when content sampling serves the dual purpose of disclosing content quality and generating advertising revenue. One of the key features of the model is that consumers evaluate free samples of their choice within the limit set by the publisher. Consumers then use the information gathered from the free samples to update their prior expectations about content quality in a Bayesian fashion to make more informed purchase decisions. Taking consumers' quality updating into account, the publisher can adopt a sampling strategy, a paid content strategy, or a free content strategy.

We derived three key results. *First*, the publisher's optimal ratio of advertising revenue to sales revenue is determined by the elasticities of expected content demand with respect to price and sample size, the price elasticity of advertising demand, and the elasticity of consumers' updated expectations with respect to the sample size. *Second*, when content quality is known to consumers, the optimal strategy is determined by the relationship between advertising effectiveness and content quality. Interestingly, it may be optimal for the publisher to offer free content samples even if sampling solely cannibalizes content demand. *Third*, when consumers learn about content quality through inspection of free samples, sampling has a demand-enhancing effect when the elasticity of consumer's posterior quality expectations with respect to sample size exceeds the ratio of sampled to paid content. In such a setting, the optimal strategy is determined by the relationship between advertising effectiveness and the interplay between quality expectations and actual content quality.

Our predictions are consistent with casual observations from the media industry (Abramson, 2010). Once advertising effectiveness is sufficiently high, our model suggests that the publisher should offer its entire content for free. Such a business model was often followed in the early days of the Internet where the provision of content was largely financed by advertising. More recently, many publishers have moved away from a pure advertising-financed business model, suggesting that either advertisers overestimated Web advertising effectiveness or that its effect has diminished over time.

Our analysis suggests several avenues for future research. *First*, regarding consumers, we assume they correctly update quality expectations based on their sample experience. One alternative is to assume a consistent bias in the consumers' judgments. In addition, in circumstances where the firm selects the samples, consumers are likely to adjust (discount) observed quality, assuming that the publisher has provided a non-representative set of samples to choose from in order to persuade them to buy the paid content. *Second*, one could assume that consumers do not evaluate the qualities of all free samples because of "sampling costs." These costs may be due to the opportunity cost of time or mental costs. *Third*, one could enrich the model by allowing for internal competition, where the publisher offers two websites to serve different categories of consumers, which relates to the versioning literature.[14] Clearly, there are many directions which research in these areas could take. We view this paper a step in this process and hope the paper encourages work in these and related directions.

References

Abramson, J. 2010. Sustaining Quality Journalism. *Daedalus*, Spring, 39–44.

Ackerberg, D.A. 2003. Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination. *International Economic Review* 44(3), 1007–1040.

Akerlof, G.A. 1970. The Market for Lemons. *Quarterly Journal of Economics* 84(3), 488–500.

Anderson, S.P., A. de Palma, and J.-F. Thisse (1992), *Discrete Choice Theory of Product Differentiation*, Cambridge, MA: MIT Press.

Anderson, S.P., J.J. Gabszewicz. 2006. The Media and Advertising: A Tale of Two-Sided Markets. *Handbook of the Economics of Art and Culture*, ed. by V. Ginsburgh and D. Throsby, Elsevier, 567–614.

Anderson, S.P., R. Renault. 2006. Advertising Content. *American Economic Review* 96, 93–113.

Bagwell, K. 2007. The Economic Analysis of Advertising. *Handbook of Industrial Organization*, ed. by M. Armstrong and R. Porter, North-Holland, vol. III, 1701–1844.

Bawa, K., R. Shoemaker. 2004. The Effects of Free Sample Promotions on Incremental Brand Sales. *Marketing Science* 23(3), 345–363.

Bhargava, H.K., V. Choudhary. 2008. Research Note: When is Versioning Optimal for Information Goods? *Management Science* 54(5), 1029–1035.

Boom, A. 2009. "Download for Free" – When Do Providers of Digital Goods Offer Free Samples? Working Paper, *Copenhagen Business School*.

[14]For instance, *The Boston Globe* operates the ad-supported site *boston.com* and the subscriber-only site *BostonGlobe.com*.

Chellappa R.K., S. Shivendu. 2005. Managing Piracy: Pricing and Sampling Strategies for Digital Experience Goods in Vertically Segmented Markets. *Information Systems Research* 16(4), 400–417.

Cheng, H.K., Q.C. Tang. 2010. Free Trial or No Free Trial: Optimal Software Product Design With Network Effects. *European Journal of Operational Research* 205, 437–447.

DeGroot, M. 1970. *Optimal Statistical Decisions*, New York: Mc Graw-Hill.

Dorfman, R., P.O. Steiner. 1954. Optimal Advertising and Optimal Quality. *American Economic Review* 44(5), 826–836.

Erdem, T., M.P. Keane, B. Sun. 2008. A Dynamic Model of Brand Choice When Price and Advertising Signal Product Quality. *Marketing Science* 27(6), 1111–1125.

Erdem, T., M.P. Keane. 1996. Decision-Making under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets. *Marketing Science* 15(1), 1–20.

Faugère C., G.K. Tayi. 2007. Designing Free Software Samples: A Game Theoretic Approach. *Information Technology Management* 8, 263–278.

Fudenberg D. and J. Tirole (1991), *Game Theory*, Cambridge, MA: MIT Press.

Gabszewicz, J.J., D. Laussel, N. Sonnac. 2004. Attitudes Toward Advertising and Price Competition in the Press Industry. *Economics of Art and Culture*, ed. by V.A. Ginsburgh, Elsevier, 61–74.

GlobeNewswire. 2012. “Wall Street Journal Offers Digital ‘Open House’ Today; Access Includes Online, Apps for *iPhone*, *iPad*.” (April 5). http://www.globenewswire.com/-newsroom/news.html?d=250593.

Godes, D., E. Ofek, M. Sarvary. 2009. Content vs. Advertising: The Impact of Competition on Media Firm Strategy. *Marketing Science* 28(1), 20–35.

Heiman, A., B. McWilliams, Z. Shen, D. Zilberman. 2001. Learning and Forgetting: Modeling Optimal Product Sampling Over Time. *Management Science* 47(4), 532–546.

Hotz, V.J., M. Xiao. 2013. Strategic Information Disclosure: The Case of Multiattribute Products With Heterogeneous Consumers. *Economic Inquiry* 51(1), 865–881

Kind, H.J., T. Nilssen, L. Sørgard. 2009. Business Models for Media Firms: Does Competition Matter for How They Raise Revenue? *Marketing Science* 28(6), 1112–1128.

Milgrom, P., J. Roberts. 1986. Price and Advertising Signals of Product Quality. *Journal of Political Economy* 94(4), 796–821.

Newspaper Association of America. 2012. Paid Digital Content Benchmarking Study.

Rysman, M. 2009. The Economics of Two-Sided Markets. *Journal of Economic Perspectives* 23(3), 125–143.

Shapiro, C., H.R. Varian. 1998. *Information Rules*, Boston, MA: Harvard Business School Press.

Sun, M. 2011. Disclosing Multiple Product Attributes. *Journal of Economics & Management Strategy* 20(1), 134–145.

Wang, C., X. Zhang. 2009. Sampling of Information Goods. *Decision Support Systems* 4(1), 14–22.

Wilson, R. 1985. Multi-Dimensional Signalling. *Economics Letters* 19(1), 17–21.

Xiang, Y., D.A. Soberman. 2011. Preview Provision Under Competition. *Marketing Science* 30(1), 149–169.

Appendix

A.1 Sampling From a Uniform Distribution

The Pareto Distribution. A random variable X has a Pareto distribution with parameters w_0 and α ($w_0 > 0$ and $\alpha > 0$) if X has a density

$$f(x|w_0,\alpha) = \begin{cases} \frac{\alpha w_0^\alpha}{x^{\alpha+1}} & \text{for } x > w_0 \\ 0 & \text{otherwise.} \end{cases}$$

For $\alpha > 1$ the expectation of X exists and it is given by $E(X) = \frac{\alpha w_0}{\alpha - 1}$. Regarding sampling from a uniform distribution, we use the following result.

Theorem (DeGroot, 1970).[15] *Suppose that $X_1,\ldots,X_n$ is a random sample from a uniform distribution of the interval $(0,W)$, where the value of W is unknown. Suppose also that the prior distribution of W is a Pareto distribution with parameters w_0 and α such that $w_0 > 0$ and $\alpha > 0$. Then the posterior distribution of W when $X_i = x_i$ ($i = 1,\ldots,n$) is a Pareto distribution with parameters w_0' and $\alpha + n$, where $w_0' = \max\{w_0, x_1,\ldots,x_n\}$.*

[15]Theorem 1, p. 172.

Proof. For $w > w_0$, the prior density function ξ of W has the following form:

$$\xi(w) \propto \frac{1}{w^{\alpha+1}}.$$

Furthermore, $\xi(w) = 0$ for $w \leq w_0$. The likelihood function $f_n(x_1, \ldots, x_n|w)$ of $X_i = x_i$ ($i = 1, \ldots, n$), when $W = w$ ($w > 0$) is given by:[16]

$$f_n(x_1, \ldots, x_n|w) = f(x_1|w) \cdots f(x_n|w) = \begin{cases} \frac{1}{w^n} & \text{for } \max\{x_1, \ldots, x_n\} < w \\ 0 & \text{otherwise.} \end{cases}$$

It follows from these relations that the posterior p.d.f. $\xi(w|x_1, \ldots, x_n)$ will be positive only for values w such that $w > w_0$ and $w > \max\{x_1, \ldots, x_n\}$. Therefore, $\xi(w|\cdot) > 0$ only if $w > w_0'$. For $w > w_0'$, it follows from Bayes' theorem that

$$\xi(w|x_1, \ldots, x_n) \propto f_n(x_1, \ldots, x_n|w)\xi(w) = \frac{1}{w^{\alpha+n+1}}$$

(the marginal joint probability density function $f_n(x_1, \ldots, x_n)$ of $X_1, \ldots, X_n$ is a normalizing constant). □

A.2 Proofs

Proof of Proposition 1. The solution to problem (2) must satisfy the first-order conditions

$$D(p, n; \tilde{V}(n)) + (p - c_s) \frac{\partial D(p, n; \tilde{V}(n))}{\partial p} + \lambda_1 = 0 \tag{A.1}$$

$$\begin{aligned} (p - c_s) \left(\frac{\partial D(p, n; \tilde{V}(n))}{\partial n} + \frac{\partial D(p, n; \tilde{V}(n))}{\partial \tilde{V}} \tilde{V}'(n) \right) & \\ + a'(n)n + a(n) + \lambda_2 - \lambda_3 &= 0 \end{aligned} \tag{A.2}$$

and the constraints $\lambda_1 p = 0$, $\lambda_2 n = 0$, and $\lambda_3(n - N) = 0$, where the λ_i's are non-negative real numbers (whose existence is assured by the Kuhn-Tucker theorem). Suppressing the arguments of content demand, (A.1) can be rewritten as

$$\frac{p - c_s}{p} = \frac{1}{\eta_p} \left(1 + \frac{\lambda_1}{D} \right). \tag{A.3}$$

Dividing (A.2) through p and substituting from (A.3) produces

$$\frac{1}{\eta_p} \left(1 + \frac{\lambda_1}{D} \right) \left(\frac{\partial D}{\partial n} + \frac{\partial D}{\partial \tilde{V}} \tilde{V}'(n) \right) + \frac{a'(n)n}{p} + \frac{a(n)}{p} + \frac{\lambda_2 - \lambda_3}{p} = 0.$$

[16]Given $W = w$, the random variables $X_1, \ldots, X_n$ are independent and identically distributed and the common probability density function of each of the random variables is $f(x_i|w)$.

Recalling that $n'(a) = \frac{1}{a'(n)}$ (from the inverse function theorem) and using the definitions of the respective elasticities, the preceding equation can be rearranged to obtain

$$\frac{pD}{an}\frac{1}{\eta_p}\left(1+\frac{\lambda_1}{D}\right)\left(\eta_n - \eta_{\tilde{V}}\tilde{V}_n\right) = \left(1-\frac{1}{\eta_a}\right)+\frac{\lambda_2-\lambda_3}{a}. \tag{A.4}$$

Under a sampling strategy there is an interior solution and hence the λ_k's are zero. Thus, (A.4) can be rewritten as

$$\frac{an}{Dp} = \frac{\eta_n - \eta_{\tilde{V}}\varepsilon_{\tilde{V}}}{(1-\frac{1}{\eta_a})\eta_p}. \qquad \square$$

Proof of Lemma 1. The optimal decisions on size of the sample and on the price follow from solving the Kuhn-Tucker conditions in Proposition 1.[17] Under a sampling strategy, the λ_k's are zero and it follows that $p^* = N\overline{V}(8(2-\phi)+\overline{V})/64$ and $n^* = N(8\phi-\overline{V})/16$. Under a paid content strategy, $\lambda_1 = \lambda_3 = 0$, leading to $p^* = N\overline{V}/4$ and $n^* = 0$. Under a free content strategy, we have that $p^* = 0$ and $n^* = N$. $\square$

Proof of Proposition 2. Using Lemma 1, it is straightforward to derive the profits under a free content strategy (FC) and a paid content strategy (PC). The profits are given by, respectively, $\pi^*_{FC} = (\phi-1)N$ and $\pi^*_{PC} = N\overline{V}/8$. Comparing the two profits shows that $\pi^*_{FC} \geq \pi^*_{PC}$ if and only if $\phi > \frac{\overline{V}}{8}+1$. The profit under a sampling strategy (SC) follows from Lemma 1 and is given by $\pi^*_{SC} = N\left(\overline{V}^2 - 16\overline{V}(\phi-2)+64\phi^2\right)/256$. Employing a sampling strategy is optimal if $\pi^*_{SC} > \pi^*_{PC}$ and $\pi^*_{SC} > \pi^*_{FC}$. It is immediate that these conditions hold if $\phi \in (\frac{\overline{V}}{8}, \frac{\overline{V}}{8}+2)$. A paid content strategy is optimal if $\pi^*_{PC} \geq \pi^*_{SC}$ and $\pi^*_{PC} \geq \pi^*_{FC}$, that is, if $\phi \leq \frac{\overline{V}}{8}$. A free content strategy is optimal if $\pi^*_{FC} \geq \pi^*_{SC}$ and $\pi^*_{FC} \geq \pi^*_{PC}$, that is, if $\phi \geq \frac{\overline{V}}{8}+2$. $\square$

Proof of Lemma 2. (a) In order to calculate $\mathbb{E}\left[\tilde{v}_0(n)\right]$ when $\bar{v}_0 < \overline{V}$, we first derive the distribution of $\tilde{v}_0(n) = \max\{\bar{v}_0, V_1, \ldots, V_n\}$. Before doing so, we state a preliminary fact: The distribution function of $M = \max\{V_1, \ldots, V_n\}$ is given by

$$\begin{aligned} F_M(t) &\equiv \Pr\{\max\{V_1,\ldots,V_n\} \leq t\} \\ &= \Pr\{\{V_1 \leq t\} \cap \ldots \cap \{V_n \leq t\}\} \\ &= \prod_{i=1}^{n} \Pr\{V_i \leq t\} = \left(\frac{t}{\overline{V}}\right)^n. \end{aligned} \tag{A.5}$$

As an immediate implication, the density function of M is given by

$$f_M(t) = \frac{nt^{n-1}}{\overline{V}^n}. \tag{A.6}$$

Next, we derive the density function of $\tilde{v}_0(n)$. By definition, $\tilde{v}_0(n)$ cannot be smaller than $\bar{v}_0$. Therefore, $\tilde{v}_0(n) = \bar{v}_0$ if and only if $\max\{V_1, \ldots, V_n\} \leq \bar{v}_0$. The probability of this event follows from (A.5) and it is given by

$$F_M(\bar{v}_0) = \left(\frac{\bar{v}_0}{\overline{V}}\right)^n.$$

[17] It is straightforward to show that the objective function is concave for all parameter values.

For $\tilde{v}_0(n) > \bar{v}_0$, let $\tilde{F}(\cdot)$ denote the truncated distribution function of $\tilde{v}_0(n)$. After removing the lower part of the distribution, we have $\tilde{F}(t) = F_M(t) - F_M(\bar{v}_0)$ for $t \in [\bar{v}_0, \overline{V}]$. This implies $\tilde{f}(t) = f_M(t)$ for $t \in [\bar{v}_0, \overline{V}]$, and hence

$$\tilde{f}(t) = \frac{nt^{n-1}}{\overline{V}^n}, \quad \text{if} \quad \bar{v}_0 \leq t \leq \overline{V}$$

by (A.6). The distribution of $\tilde{v}_0(n)$ has a mixed structure with

$$\Pr\{\tilde{v}_0(n) = \bar{v}_0\} = \left(\frac{\bar{v}_0}{\overline{V}}\right)^n \tag{A.7}$$

and density

$$\tilde{f}(t) = \frac{nt^{n-1}}{\overline{V}^n}, \quad \text{if} \quad \bar{v}_0 \leq t \leq \overline{V}. \tag{A.8}$$

The expectation of this mixed distribution is given by

$$\begin{aligned} \mathbb{E}[\tilde{v}_0(n)] &= \bar{v}_0 \left(\frac{\bar{v}_0}{\overline{V}}\right)^n + \int_{\bar{v}_0}^{\overline{V}} \frac{nt^n}{\overline{V}} dt \\ &= \frac{\bar{v}_0^{n+1} + n\overline{V}^{n+1}}{(n+1)\overline{V}^n}. \end{aligned}$$

Substituting this expression into (11) produces (12). (b) If $\bar{v}_0 \geq \overline{V}$, then $\tilde{v}_0(n)$ is equal to $\bar{v}_0$, which in turn implies that $\mathbb{E}[\tilde{v}_0(n)] = \bar{v}_0$. Substituting this expression into (11) yields (13). □

Proof of Lemma 3. If $\bar{v}_0 < \overline{V}$, the quality gap can be expressed as

$$\tilde{V}(n) - \frac{\overline{V}}{2} = \frac{\bar{v}_0^{n+1}(\alpha+n) - \overline{V}^{n+1}(\alpha-1)}{2(\alpha+n-1)(n+1)\overline{V}^n}. \tag{A.9}$$

Clearly, the sign of the quality gap depends only on the sign of numerator (A.9). The latter can easily be rearranged to obtain (16). If $\bar{v}_0 \geq \overline{V}$, the quality gap can be written as

$$\tilde{V}(n) - \frac{\overline{V}}{2} = \frac{(\alpha+n)\left(\bar{v}_0 - \overline{V}\right) + \overline{V}}{2(\alpha+n-1)},$$

which is strictly positive by our assumptions. □

Proof of Lemma 4. Differentiating (17) with respect to n yields

$$\frac{\partial D^E(p,n)}{\partial n} = \frac{\left((N-n)\tilde{V}(n)' - \tilde{V}(n)\right)p}{\left((N-n)\tilde{V}(n)\right)^2}.$$

Clearly, sampling is demand-enhancing if $(N-n)\tilde{V}'(n) - \tilde{V}(n) > 0$, which can be rewritten as $\frac{\tilde{V}'(n)n}{\tilde{V}(n)} > \frac{n}{N-n}$. □

Proof of Proposition 3. At an interior solution, the optimal sample size n^* satisfies the first-order condition

$$(N-n^*)\frac{\tilde{V}'(n^*)}{4}-\frac{\tilde{V}(n^*)}{4}+\phi-\frac{2n^*}{N}=0.$$

For a corner solution involving $n^*=0$, the Kuhn-Tucker conditions imply

$$\frac{N\tilde{V}'(0)}{4}-\frac{\tilde{V}(0)}{4}+\phi\leq 0 \quad\Longleftrightarrow\quad \phi\leq\underline{\phi}.$$

At the other extreme, when $n^*=N$, the Kuhn-Tucker conditions require that

$$-\frac{\tilde{V}(N)}{4}+\phi-2\geq 0 \quad\Longleftrightarrow\quad \phi\geq\overline{\phi}.$$

□

Proof of Lemma 5. Using the definition of $\tilde{V}(n)$ in (15), the lower bound can be expressed in terms of the underlying model parameters as

$$\underline{\phi}=\frac{(2\alpha(\alpha-1)+N)\bar{v}_0}{16(\alpha-1)^2}.$$

Setting $\bar{v}_0=\overline{V}$ and letting $\alpha\to\infty$ yields that $\underline{\phi}\to\frac{\overline{V}}{8}$. Likewise, we have that

$$\overline{\phi}=\frac{(\alpha+N)\overline{V}}{8(\alpha+N-1)}+2.$$

Letting $\alpha\to\infty$, we obtain $\overline{\phi}\to\frac{\overline{V}}{8}+2$. □